

A PRACTICAL OSINT GUIDE TO CONDUCTING BUSINESS INVESTIGATIONS AND BACKGROUND VERIFICATIONS

OPEN-SOURCE INTELLIGENCE TECHNIQUES FOR
SMART BUSINESS



A PRACTICAL OSINT GUIDE TO CONDUCTING BUSINESS INVESTIGATIONS AND BACKGROUND VERIFICATIONS

**OPEN-SOURCE INTELLIGENCE TECHNIQUES
FOR SMART BUSINESS**

INDIA

© Surana & Surana International Attorneys.

Table of Contents

Open-source intelligence techniques for smart business	1
Introduction	4
Applications of Open-Source Intelligence.....	6
The roadmap.....	7
Tools and techniques	10
Global business database.....	10
Zaubacorp	10
Trademarks search	14
Verify DIN/DPIN-PAN and Aadhar details of Director/ Desgnated Partner.....	15
Verify court records	18
e-Courts services.....	18
Legitquest.....	20
News/ Articles/ Blogs.....	22
Social Mention	22
Leaked databases.....	23
Data leaks.....	24
Haveibeenpwned?	25
Search engines	26
DuckDuckGo.....	26
Carrot2	26
TOR browser	27
Website investigation	27
Domain search	27
Web-archives	29
Email investigation.....	30
Automated tools	30
Maltego.....	30
Skopenow.....	31
Cobwebs.....	31
Conclusion.....	31

List of figures

Figure 1: Example demonstrating an investigation plan for business entities.....	9
Figure 2: Open corporates search query results for Credit-Suisse India located in Pune.....	10
Figure 3: Open corporates search query results for Credit-Suisse branches.	10
Figure 4: Zaubacorp search options.....	11
Figure 5: Example demonstrating functioning of Zaubacorp.com.	11
Figure 6: Company details from Zaubacorp.....	12
Figure 7: List of financial documents (paid) that may be available for some business listings on Zaubacorp.	13
Figure 8: Directorship information on Zaubacorp.	14
Figure 9: Public search of Trademarks, GOI portal.	15
Figure 10: View master data about a company and its designated partner on IP India website.	15
Figure 11: DIN lookup for designated director will open another box asking for full name, father's last name and date of birth on IP India website- master data search.....	16
Figure 12: Online PAN/ TAN verification.	17
Figure 13: Verification of Aadhaar number online on UIDAI portal.	18
Figure 14: e-Courts of India portal to search for past legal records with CNR number or petitioner/ respondent's name.	19
Table 1: Contains a list of other portals from where case data can be accessed.....	18
Figure 15: Search results on Legitquest.com with case name or with the title.	21
Figure 16: Real time news search on Social Mention.	22
Figure 17: Search results from an Offshore leaked database filtered to the Indian jurisdiction.	23
Figure 18: Database containing investigation maps that link people with the entities and the positions held by them (<i>The International Consortium of Investigative Journalists, 2021</i>)....	24
Figure 19: haveibeenpwned results for a demo email id. It has been found in 8 data breaches.	25
Figure 20: List of breaches in which the demo email id was found.	26
Figure 21: Whois example for domain name facebook.com.....	27
Figure 22: Registrant details for domain name facebook.com.	28
Figure 23: Whois details about administration and technical contact for domain name facebook.com.....	29

Disclaimer

The handbook titled "*A Practical OSINT Guide to Conducting Business Investigations and Background Verifications*" is intended to assist businesses ranging from individual-led ones to corporates. Most of the resources available are focused on the information available for the US, UK, and European countries. This guide is an attempt to portray OSINT from an Indian perspective and the resources available in India. The information within the handbook contains real information, however if the tools and the portals change, the author bears no responsibility for futuristic errors.

About the author

Dr. Vinod Surana holds a Ph.D in International Telecommunication regulations from Madras University and a Masters in Law from Cornell University. He has undergone executive programs at Indian Institute of Management (Ahmedabad), Indian Institute of Management (Bangalore), Indian School of Business (Hyderabad) as well as been sponsored by the Governments of USA, Germany and Japan to undergo specialized professional and management training programs in respective countries. He is the CEO of Surana and Surana International Attorneys, Chennai and a sought after speaker, mentor and writer on various subjects including Law, Defence, Technology, International Affairs, Alternate healing etc.

INTRODUCTION

In this day and age, the internet is the most accessible and convenient way to publish and retrieve information. Modern investigators utilize resources on the web to navigate through their investigations. The secret lies within the core of unclassified data especially with smart devices at our disposal. Posting about our digital lives has become a ubiquitous trend and digital marketing solutions have made businesses of all scales visible and accessible from anywhere in the world.

One of the most sought-after ways of choosing a restaurant to eat at is by perusing its reviews on Google. Similarly, when hiring an employee, companies scrutinize future employees' online activities to gain an insight into their thought process (Edward J Apple, 2011; Mehta, 2020a). Blacklisted individuals can be filtered with an OSINT approach. Social media investigations have been excluded from the scope of this guide.

APPLICATIONS OF OPEN-SOURCE INTELLIGENCE

OSINT (Open Source Intelligence) can be defined as data collected from publicly available sources transformed into actionable intelligence. OSINT is the link between people, entities, social and personal events. Open-source information is termed as OSINF. Nearly, 57% of the total world's population uses social media equating to around 4.8 billion as of July 2021 (Data Reportal, 2021). Out of which, over 340 million Indians use Facebook making India a leading country in the growth of their platform (Statista, 2021a). India ranks third in the world in terms of Twitter users with 22.10 million users (Statista, 2021b).

Digital adoption continues to rapidly grow in India. Technological advancements in the field of artificial intelligence, machine learning and the advent of smart devices have only propelled the amount of information available in the cyber space. Additionally, India is the second largest online market in the world after China, implying that the number of internet users and online businesses are only likely to grow. Remote or work from home flexibility have taken companies into complex situations where they are struggling to maintain privacy and avoid potential theft of sensitive information. More often, work from home employees take up secondary employments owing to the comforts of their homes and the flexibility in work schedules to earn that extra penny.

Remote working also increases threats to cyber-attacks like Phishing. Phishing is a type of social engineering attack that steals user credentials (Jahankhani, Al-Nemrat and Hosseinian-Far, 2014; Rosenthal, 2020). For example, an employee receives an email from a 'trusted entity' with the intention of tricking him into revealing sensitive information. The source may appear genuine, but the attacker has impersonated someone in order to lure the user in. When the employee clicks on the link, he is redirected to an authentic 'looking' page that prompts him to enter his credentials.

Alternately, if he were to download the attachment in the email, a payload will be automatically installed in the employee's system thereby compromising the data- both personal and professional. In such scenarios, it is vital for business entities to train their employees, increase cyber awareness about on-going threats and modus operandi. There are several practical implications of OSINT investigations.

The law enforcement community applies OSINT techniques to gather information about their targets in criminal and civil procedures (Bazell, 2019; Pastor-Galindo et al., 2020). On the other hand, private corporate companies use OSINT to conduct background checks on future employees, present employees, management, clients, and consumers. Detectives observe the social media presence of their targets and gather information about the target's inner circle of friends and families to establish crucial links between them (Bazell, 2019; Samantha Elizabeth Rule, 2014).

Photographs and video footages are closely monitored to identify possible locations and time when it was shot (Furuhaug, 2019; Reuser, 2017). The pandemic induced financial stress could potentially cause an increase in corporate theft. Money laundering, frauds committed by business entities as well as deceitful encounters within the corporate realm can come across as a nuisance (Button and Cross, 2017; Infosys, 2020). Conducting background verification checks on individuals, business entities or deep diving into the sea of information on social media can help safeguard businesses in various ways (Bazell, 2019; Edward J Apple, 2011).

This guide will attempt to answer some of the specific questions related to conducting **OSINT in India** and also touch upon certain tools that can be useful for the international community as well.

THE ROADMAP

Imagine a situation where your company is working hard on creating a lucrative software for the government that is geared towards promote tourism. The CEO, Mr. A receives a business proposal by another company 'X' aiming towards a similar invention. The proposal requests for a partnership between the firms and the idea seems achievable. The project also involves hiring a team of eight members allocated to various roles as an immediate requirement. Mr. A wants to have a clear understanding of company X and the people working there. You are given the responsibility to conduct OSINT on X and report to the CEO.

The first step in every investigation is to *prepare*. The aim is to answer the following questions while creating an investigation map (Mehta, 2020b):

1. What is already known?- name of the company, owner/ director, location etc.
2. What is the aim of the investigation?- Nature of business, status, background check etc.
3. What needs to be found out?- Should your company partner with X?
4. How can the missing attributes be searched for?- Locate X on tourism posts, websites, government sites etc.
5. Where is the investigation leading?- Monitor at every step.

6. How can one attribute link to the other and that with another and so on?- Mind maps (*figure 1*) help!

The investigation roadmap may look like (*figure 1*). As the search progresses, more and more attributes start populating the map and the findings often demonstrate a bigger picture.

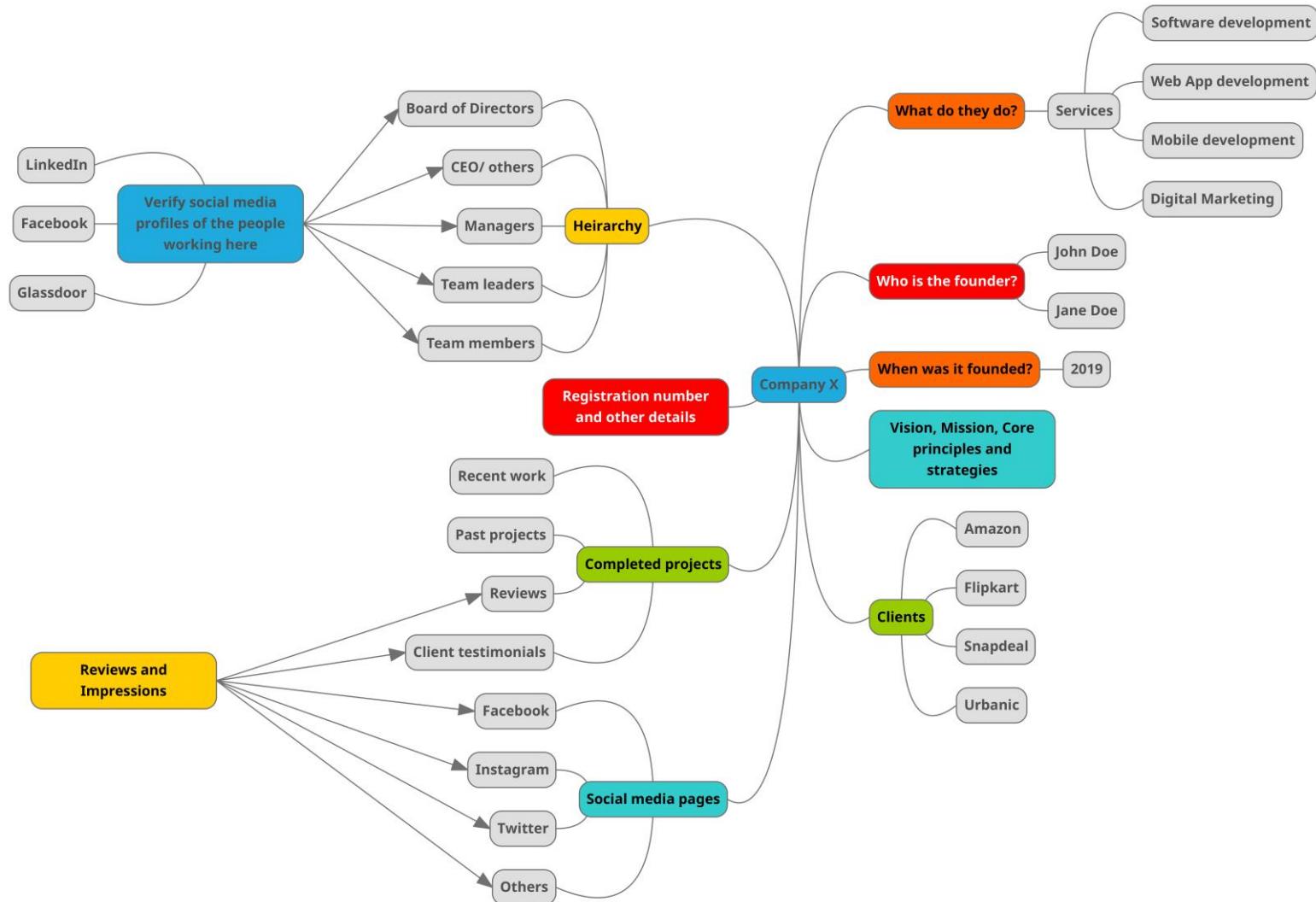


Figure 1: Example demonstrating an investigation plan for business entities.

TOOLS AND TECHNIQUES

For demonstration purposes, company and individual names have been chosen at random to demonstrate how OSINT tools yield results.

GLOBAL BUSINESS DATABASE

www.Opencorporates.com contains the largest open database of companies worldwide (1,98,706,341 listed as of August 2021). Search queries can be filtered based on jurisdiction, name of the company, name of the officers and the company number. The results include registered company number (along with industry code), information as to the company's status whether, it is active or inactive, type of company, registered address, listed directors and officers (*figure 2*). It also includes a timeline of events and company network that details the time of removal or addition of an officer.

CREDIT SUISSE SERVICES (INDIA) PRIVATE LIMITED

Company Number	U93090PN2006F	69
Status	Active	
Incorporation Date	3 October 2006 (almost 15 years ago)	
Company Type	Company limited by Shares	
Jurisdiction	India	
Registered Address	Ground Floor, Wing 1, Cluster A, EON FREE ZONE Plot No.1, S. No. 77, MIDC Knowledge Park, Kharadi Pune Pune MH 411014 IN India	
Industry Codes	93090: (India National Industrial Classification 2004 (MCA 2009))	
Directors / Officers	GIRISH MEHRA, 5 Dec 2018- JOHN BURNS, 6 Mar 2017- RAJIV RAMACHANDRAN, 22 May 2018- ZAHABIYA HUSSAIN OFFICEWALA, 18 Sep 2019-	

Company network

Not yet available for this company. [Click to find out more](#)

Latest Events

- 2018-12-05 Addition of officer GIRISH MEHRA,
- 2019-06-24 - 2019-10-30 Removal of officer RANJIT KUMAR ANAND,
- 2019-09-18 Addition of officer ZAHABIYA HUSSAIN OFFICEWALA,

[See all events](#)

Corporate Grouping USER CONTRIBUTED

None known. [Add one now?](#)

[See all corporate groupings](#)

Similarly named companies

[View all similarly named companies](#)

Furthermore, it collates associated branches linking their details (*figure 3*).

Branches •

 inactive branch CREDIT SUISSE SERVICES (INDIA) PRIVATE LIMITED (United Kingdom, 16 Mar 2015-31 May 2019) details
 inactive branch CREDIT SUISSE SERVICES (INDIA) PRIVATE LIMITED (United Kingdom, 16 Mar 2015-31 May 2019) details

Figure 3: Open corporates search query results for Credit-Suisse branches.

ZAUBACORP

www.Zaubacorp.com assists in finding financial information of businesses and provides access to critical documents required for fact checking purposes. Documents related to appointment and resignation of Directors, incorporation of business, forms filed with Registrar of companies and lots more is available on this website. It is one of the leading providers of commercial information on businesses in India. Search can be conducted using keywords related to the company, director, trademark, or address if known (figure 4).

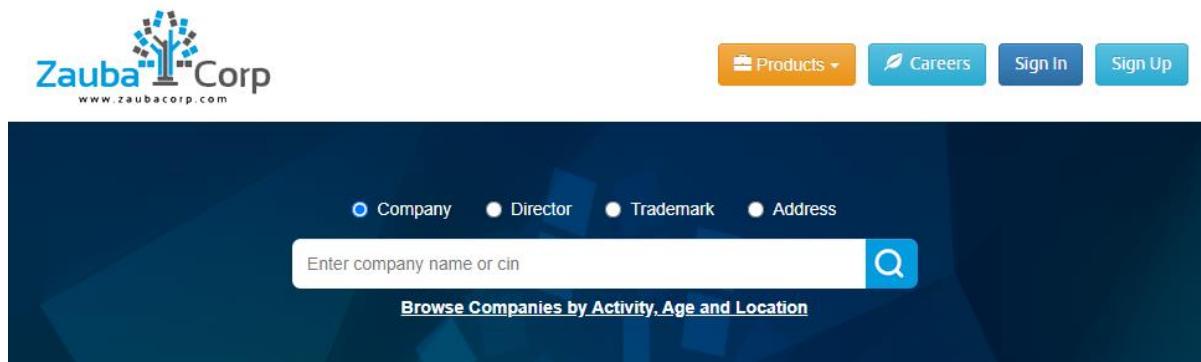


Figure 4: Zauba Corp search options

Figures 5 and 6 illustrate basic information and details about the registered company.

A screenshot of the Zauba Corp company profile for 'FARO INDIA PRIVATE LIMITED'. The company name is at the top left, with a date 'As on: July 16, 2021'. To the right is a 'Track this company' button. Below the company name are six circular icons with labels: 'Basic Information' (building icon), 'Documents' (document icon), 'Trademarks' (circle with 'R' icon), 'Directors' (person icon), and 'Map' (location pin icon). The 'Basic Information' section contains the following details: 'Faro India Private Limited is a Private incorporated on 09 December 2004. It is classified as Non-govt company and is registered at Registrar of Companies, Delhi. Its authorized share capital is Rs. and its paid up capital is Rs. It is involved in Wholesale of machinery, equipment and supplies.' Below this is a note: 'Faro India Private Limited's Annual General Meeting (AGM) was last held on 31 December 2020 and as per records from Ministry of Corporate Affairs (MCA), its balance sheet was last filed on 31 March 2020.' The 'Directors' section lists 'Directors of Faro India Private Limited are:'. The 'Map' section shows a small map of India with a pin. The 'Current status of Faro India Private Limited is - Active.'

Figure 5: Example demonstrating functioning of Zauba Corp.com.

Company Details

CIN	53
Company Name	FARO PRIVATE LIMITED
Company Status	Active
RoC	RoC-Delhi
Registration Number	3
Company Category	Company limited by Shares
Company Sub Category	Non-govt company
Class of Company	Private
Date of Incorporation	09 December 2004
Age of Company	16 years, 8 month, 11 days
Activity	Wholesale of machinery, equipment and supplies Click here to see other companies involved in same activity.

Figure 6: Company details from Zaubacorp.

Figure 7 enlists finance related documents that may be available for some business listings on Zaubacorp.

Financial Report

Balance Sheet

Paid-up Capital	🔒
Reserves & Surplus	🔒
Long Term Borrowings	🔒
Short Term Borrowings	🔒
Trade Payables	🔒
Current Investments	🔒
Inventories	🔒
Trade Receivables	🔒
Cash and Bank Balances	🔒

Profit & Loss

Total Revenue (Turnover)	🔒
Total Expenses	🔒
Employee Benefit Expenses	🔒
Finance Costs	🔒
Depreciation	🔒
Profit Before Tax	🔒
Profit After Tax	🔒

Figure 7: List of financial documents (paid) that may be available for some business listings on Zaubacorp.

Additionally, list of current directors, appointment and cessation dates are also enclosed within the search results. One can easily navigate through other companies under the same directorship to find substantial information. The database also runs search parameters based on companies having similar addresses that can further elaborate key evidence.

Director Details

DIN	Director Name	Designation	Appointment Date	
0 4	AI 'GI	Director	16 October 2020	View other directorships
08661691	AL H	Director	08 January 2020	View other directorships
08806159	MA I	Director	23 July 2020	View other directorships

Past Director Details

DIN	Director Name	Appointment Date	Cessation Date	
02043630	KI R	16 October 2006	01 April 2015	View other directorships
03431289	JO NE	15 March 2011	06 April 2018	View other directorships
06922183	N	07 August 2015	20 December 2018	View other directorships
02043695	JA	16 October 2006	04 December 2015	View other directorships
07291671	MI	16 June 2014	07 August 2015	View other directorships
08104374	ROI	06 April 2018	31 January 2020	View other directorships
05189651	SUI	20 December 2018	12 March 2021	View other directorships

Figure 8: Directorship information on Zaubacorp.

Tip: It will be interesting to find what information is available on Opencorporates and Zaubacorp about your own business/organization.

TRADEMARKS SEARCH

The website of Ministry of Corporate Affairs provides a list of services on their portal at <https://www.mca.gov.in/mcafoportal/checkCompanyName.do>. It enables the user to check company name, trademark, and domain of the proposed name. Sites like these would be useful to run simple searches to see whether the name is available and if it has been trademarked or not.

Another portal for public search of Trademarks is <https://ipindiaonline.gov.in/tmrpublicsearch> (figure 9).



Class Details | Well Known Marks | Prohibited Marks | Vienna Code Classification | International Non-Proprietary Names(INN) | Help

Search Type : Wordmark

Key Word	Value
Wordmark	<input type="text"/> Start With
Class	<input type="text"/> *Enter one class at a time
Goods Description	<input type="text"/> *Will be available soon

Search

Figure 9: Public search of Trademarks, GOI portal.

VERIFY DIN/DPIN-PAN AND AADHAR DETAILS OF DIRECTOR/ DESGINATED PARTNER

IP India website allows users to verify DIN/DPIN details of Designated partners <https://www.mca.gov.in/mcafoportal/verifyDIN.do>. However, one must know the identification numbers and income tax permanent account number to conduct a search. The portal also allows to view master data about a company and its respective director(s).

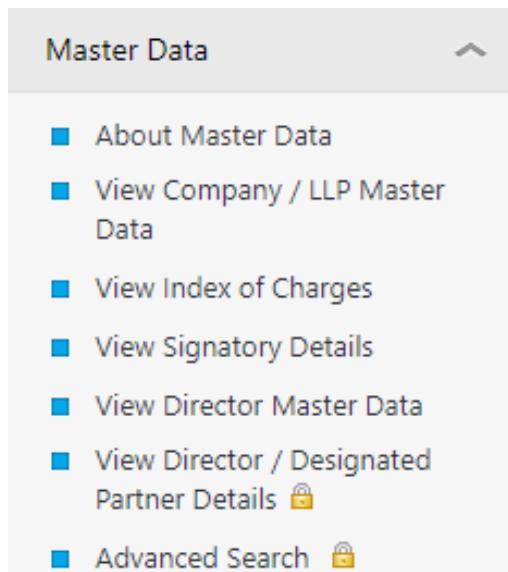


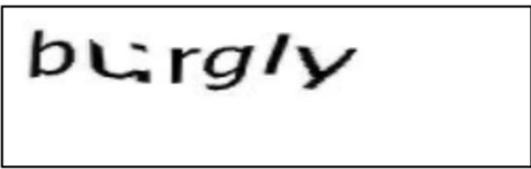
Figure 10: View master data about a company and its designated partner on IP India website.

However, DIN lookup for designated partner will require full name of the director, father's last name and date of birth of the director (figure 11).

View Director Master Data

Director Name 

DIN*

Enter Characters shown below : 


DIN Lookup

Director/ Designated Partner Name *

Father's Last name

Date of Birth  (dd/mm/yyyy)

Figure 11: DIN lookup for designated director will open another box asking for full name, father's last name and date of birth on IP India website- master data search.

Online PAN and TAN verification for India can be done on <https://incometaxindiaefiling.gov.in/>

Steps to Verify PAN details

Step-1

Logon to 'e-Filing' Portal www.incometaxindiaefiling.gov.in

Step-2

Click on 'Verify Your PAN details' hyperlink from the 'Quick Links' Section.

Step-3

Enter the PAN, Full Name (As per PAN), Date of Birth and Choose the 'Status' as applicable.

Step-4

Enter the Captcha as in the image and click on 'Submit' to verify your PAN details.

Once details including PAN number, Full name, Date of birth and mobile number are entered- an OTP will be sent to the registered number for verification.

Caution: The PAN holder will be notified with an OTP; hence permissions must be taken to verify PAN.

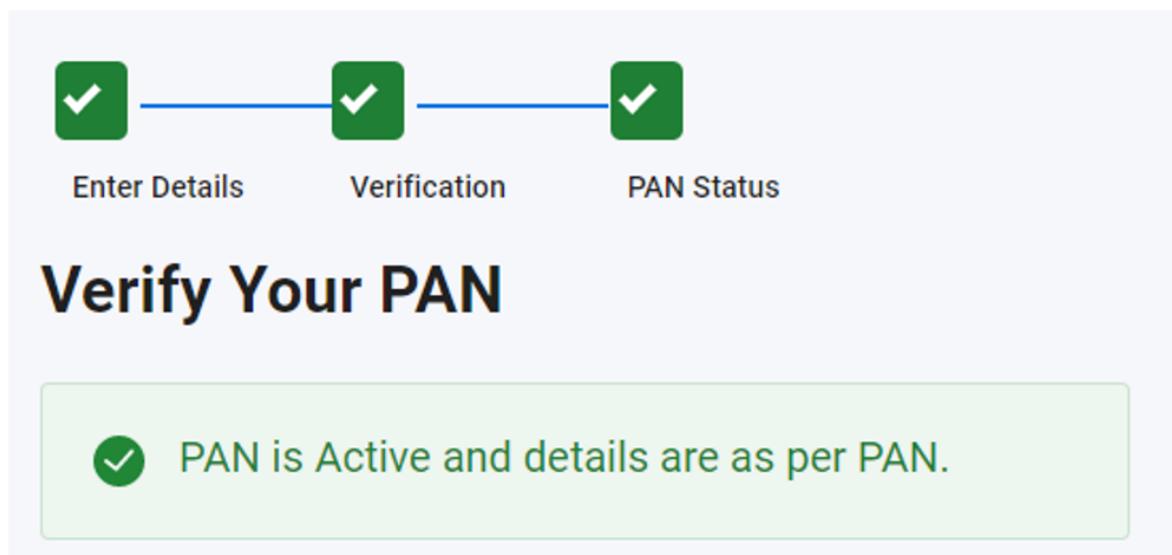


Figure 12: Online PAN/ TAN verification.

Verification of Aadhaar number can also be done online at <https://resident.uidai.gov.in/verify>.

Verify Aadhaar

Here you can check if your Aadhaar or Aadhaar submitted to you is a genuine one or not. Resident's are using this service to verify the identity of their workers.

Enter 12 digit Aadhaar number (UID)

Aadhaar Number * 

12 Digit UID (1234/1234/1234)

Captcha Verification * 

Type the character you see in the picture.




After entering the 12 digit Aadhaar number, the portal verifies it with a green tick as shown below.

Enter 12 digit Aadhaar number (UID)

Aadhaar Number * 

12 Digit UID (1234/1234/1234)

4	8	2	1
----------	----------	----------	----------

Figure 13: Verification of Aadhaar number online on UIDAI portal.

VERIFY COURT RECORDS

A criminal record check is essential for not only business-related background checks but also for matters related to our personal lives. One can verify court records to ensure that the company is not hiring someone who has been prosecuted for fraud, embezzlement or grievous crimes like rape or murder. It is imperative for the wellbeing and reputation of the company to verify their business associates and people working in the establishment. There are many databases where one can research on an applicant's past criminal records, such as:

E-COURTS SERVICES

The portal allows users to search for cases with CNR number or with the name of the petitioner/respondent. It includes databases from the District and Taluka courts of India (figure 14).

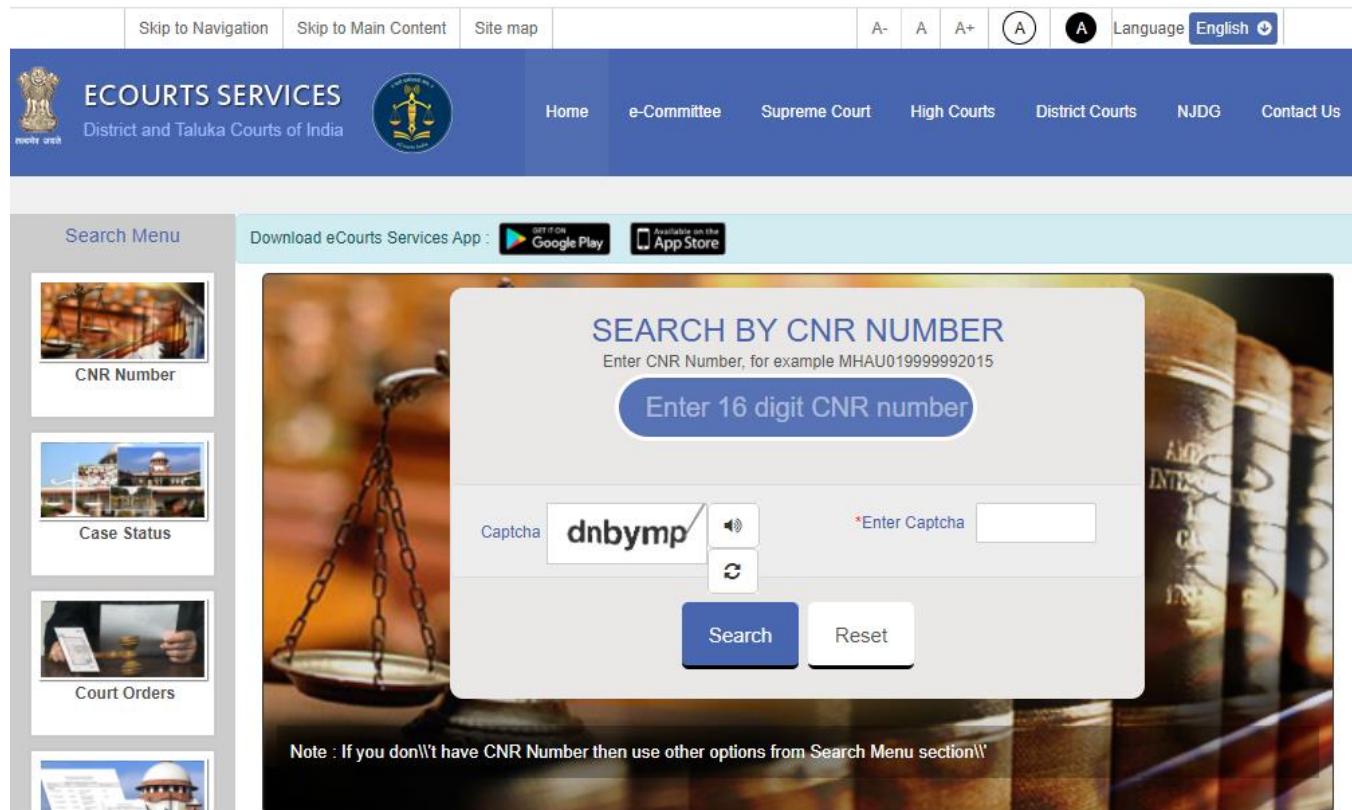


Figure 14: e-Courts of India portal to search for past legal records with CNR number or petitioner/respondent's name.

There are 39 High Court complexes in India. As of 20th August 2020, 1.5 million High Court cases were listed with 5.77 million pending cases and 31.4 million disposed cases. Similarly, there are 3296 District and Taluka complexes in India. 771.28K cases were listed on this day with 40.12 pending cases and 1.68 disposed cases. To conclude, the database is regularly updated for hearings all over the country. Table 1 contains other resources for searching legal/criminal records.

Table 1: Contains a list of other portals from where case data can be accessed.

Court	Portal	Comments
Supreme Court of India	https://main.sci.gov.in/	Enables users to view case status with diary number/ case number/ party number etc. Past judgements can be accessed by inputting similar information as above. Advance filters like name of the Judge, parties, Acts etc can help to funnel the search.
High Court	Portals are available for every individual High Court in India. For example: Madras High Court http://www.hcmadras.tn.nic.in	Cases can be retrieved through case number, title of the petitioner or respondent and/ or advocate name.

LEGITQUEST

www.legitquest.com is a structured legal database with features like one-click judgement evaluation systems. It contains more than 50 million pages of case laws of Indian database from all courts, news, interviews, and columns across the board. Legitquest relies on deep technology, artificial intelligence and neural networks that help in retrieving data quickly. *Figure 15* is an example of the search results simply by entering keywords, case name or title.

The screenshot shows the Legitquest website interface. At the top, there is a search bar with the query "Kingfisher Airlines Limited v. State Bank" and a magnifying glass icon. To the right of the search bar are buttons for "All" (selected), "Sign In", and "Sign Up". Below the search bar, a search filter section is visible with the heading "SEARCH FILTER" and a "Clear" button. The filter includes a "Filter Within Result" button and dropdown menus for "By Court" (Supreme Court, High Court, Tribunal), "By Bench", "By Year", and "By Dispositions". To the right of the filter, search results are displayed. The results page shows a summary of 47 results for the query, with a "Sort by Relevance" dropdown. Below this, a list of judgments is shown, with the first result being "Kingfisher Airlines Limited V. Union Of India" from the "High Court Of Judicature At Calcutta" on 26-09-2014. The result includes a link to the full judgment, a "Search Within Case" button, and a "Judges" button. Below this, another result is shown: "Kingfisher Airlines Limited V. Union Of India Through The Secretary, Ministry Of Finance, Government Of India..." from the "High Court Of Judicature At Bombay" on 15-07-2015.

After clicking on the URL- it redirects to the details of the case.

The screenshot shows the detailed view of the case "Kingfisher Airlines Limited v. State Bank Of India & Others". The top navigation bar includes "iDraf" and dropdown arrows. Below the navigation, there are tabs for "Judgment" (selected), "Future Reference", "Cited In", "Advocates", "Bench", and "Eq Citations". The main content area displays the case title "Kingfisher Airlines Limited V. State Bank Of India & Others" and the court "High Court Of Karnataka". The case number is "Original Side Appeal No. 1 Of 2014 | 29-01-2014". The decision is titled "Kingfisher Airlines Limited V. State Bank Of India & Others" and is dated "29-01-2014". The decision is attributed to "High Court Of Karnataka".

Figure 15: Search results on Legitquest.com with case name or with the title.

Other websites such as Indian Kanoon, LexisNexis, Law Finder Live can be used in a similar context.

NEWS/ ARTICLES/ BLOGS

Sometimes it may be important to stay up-to date with past and current happenings. It may be vital to monitor competitor activities and possess updated information about their current activities. OSINT tools like Talkwalker (www.talkwalker.com), Social mention (www.socialmention.com) etc. help to monitor social media.

SOCIAL MENTION

Social mention allows users to run real-time searches that include blogs, microblogs, bookmarks, images, and videos. Advanced search preferences enable users to choose specific dates to customize results. Additionally, frequently used keywords are pulled out from the news (figure 16).

Tip: Remember to use these keywords to conduct social media search, hashtags can be very useful to find relevant content

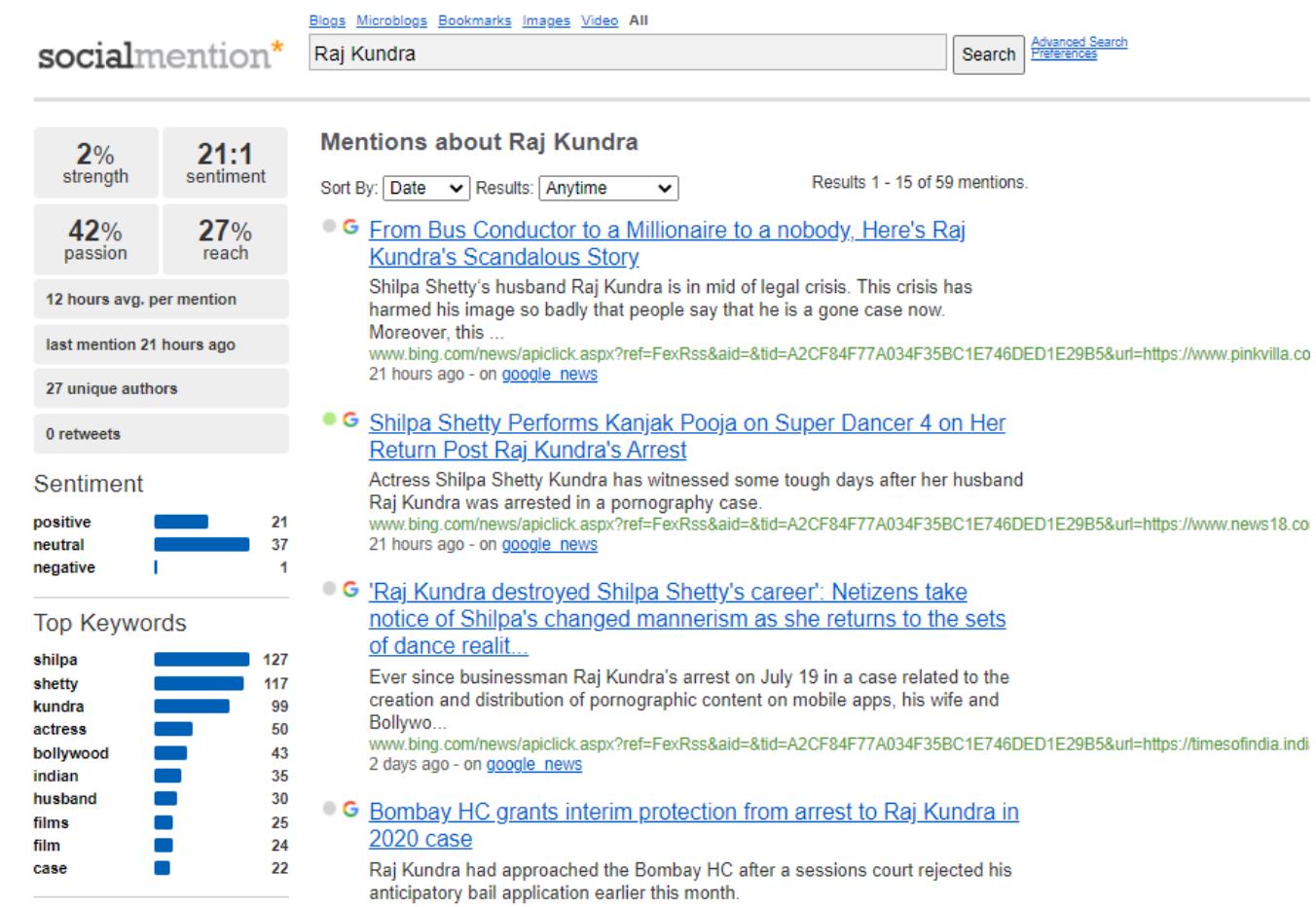


Figure 16: Real time news search on Social Mention.

Data from social mention can be downloaded in csv files for public sentiment, top keywords, top users, and top hashtags.

Other tools like Sprout social, Zoho social and Hootsuite are recommended for the same as well.

LEAKED DATABASES

ICIJ (The International Consortium of Investigative Journalists) offshore leaks contains information of more than 7,85,000 offshore entities that are a part of the Paradise Papers (2017-2018), the Panama Papers (2016), the Offshore Leaks (2013) and the Bahama Leaks investigations (2016). The information contained in this database helps to link people to companies in more than 200 countries and territories (The International Consortium of Investigative Journalists, 2021). It is a very useful database for finding data about tax havens and exposes the real names of the owners of such properties (*figure 17*).

OFFSHORE LEAKS DATABASE

DONATE NOW

Enter a name here or just filter by country to see a full list of nam

Country India

SEARCH

VENUS HOLDING CORPORATION

LYNX HOLDINGS INC.

MIDLAND INTERNATIONAL INVESTMENTS LTD.

Millenium Group Investments Ltd.

Panama Papers

Panama Papers

Panama Papers

Panama Papers

Figure 17: Search results from an Offshore leaked database filtered to the Indian jurisdiction.

The data includes a map illustrating links between people and their positions (figure 18).

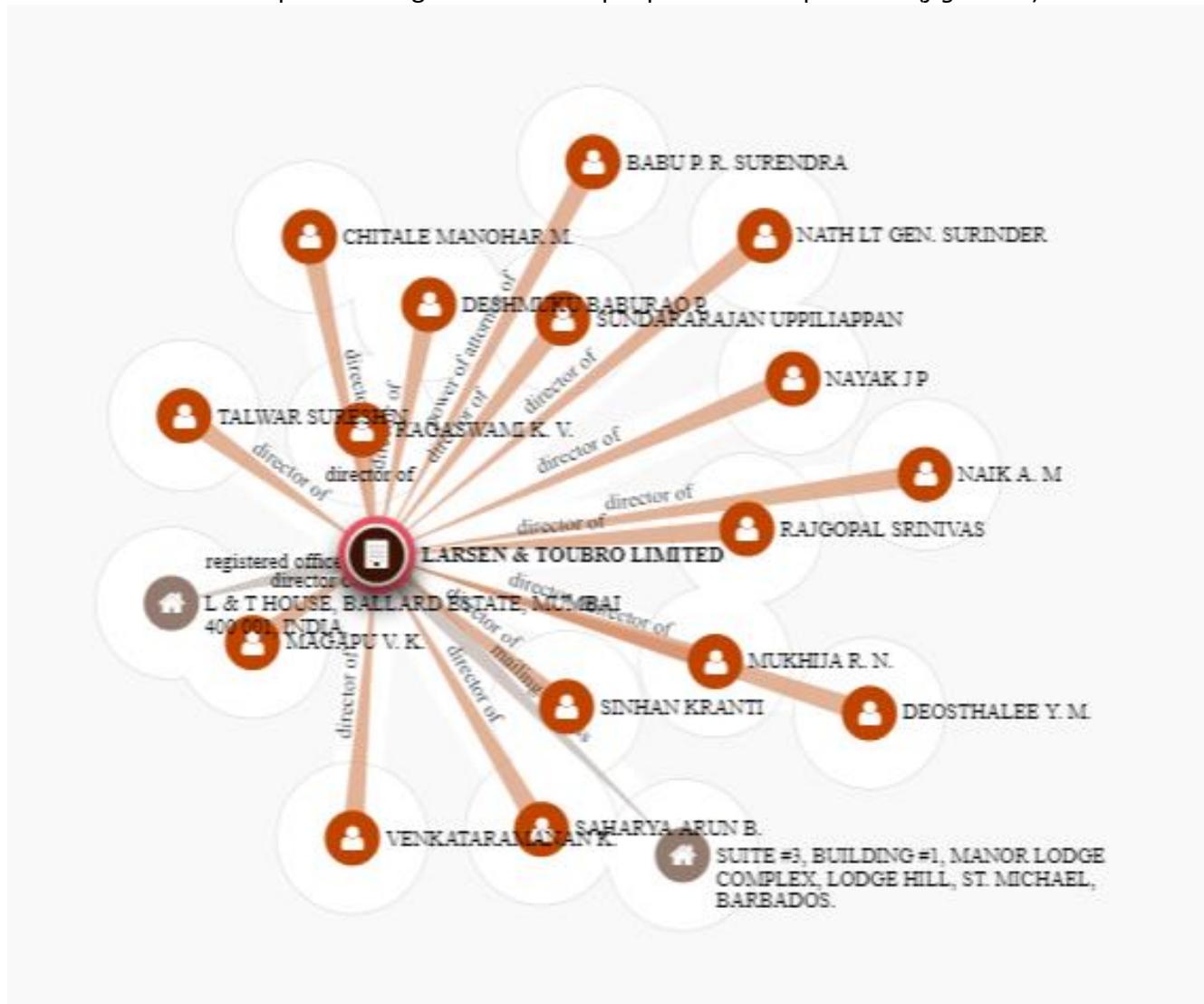


Figure 18: Database containing investigation maps that link people with the entities and the positions held by them
(*The International Consortium of Investigative Journalists, 2021*).

Alongside, information related to the address with pin code, link to the registry (Panama/Paradise/Offshore leaks/Bahama Papers) and connections of single entities are also provided on the portal.

DATA LEAKS

Data leaks and data breaches happen very frequently. Massive data breaches like that of Facebook, Alibaba, Target, Yahoo or Canva exposed confidential information about its users by unauthorized entities. Data breaches often include details like full name, date of birth, credit card/ debit card credentials, secret questions, usernames, and passwords. For business entities, it is essential to probe regularly into the nature of such data belonging to their organization. It may include official email ids, usernames and passwords of the employees, confidential information stolen from the office and pasted on sites like Pastebin etc.

Data leaks may occur because of ransomware or cyber-attacks too where the intention is to expose or defame an institution or an individual. Whistle-blowers from the company can leak information

that may be useful to competitors. Business entities can analyse such information to gain more knowledge about their competitors.

HAVE I BEEN PWNED?

It contains a database of billions of leaked credentials belonging to compromised accounts in the events of data breaches/leaks (*figure 19*).

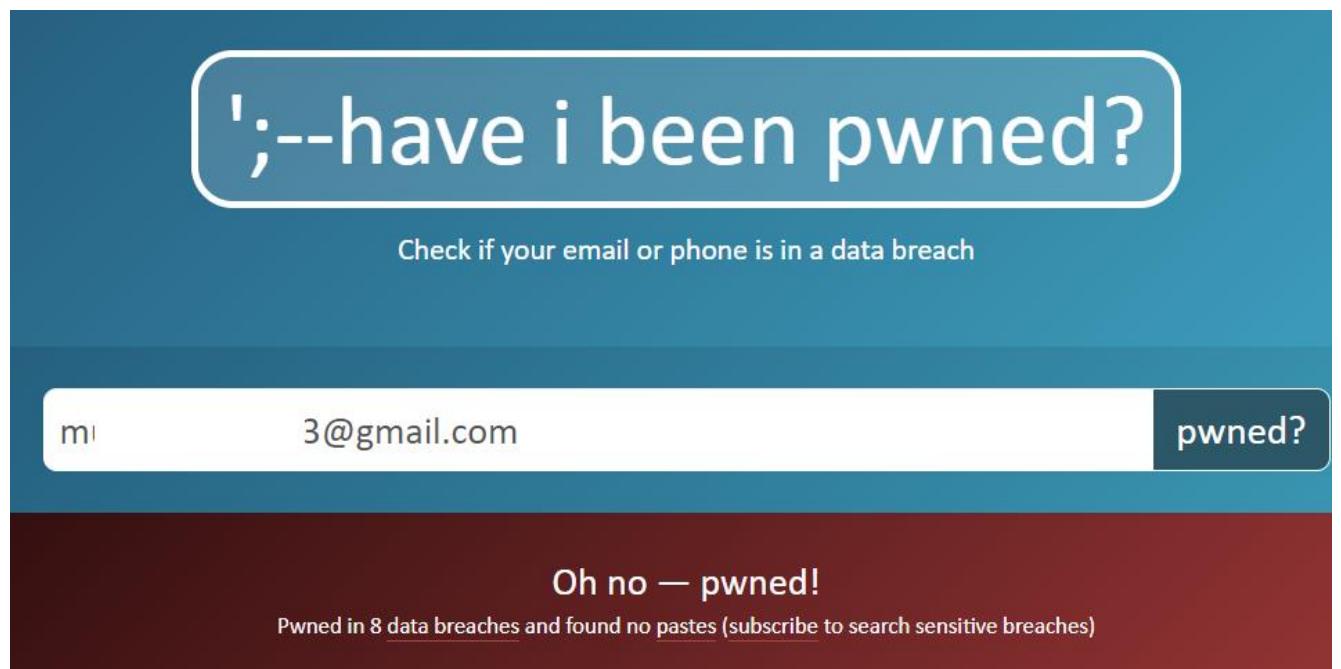


Figure 19: haveibeenpwned results for a demo email id. It has been found in 8 data breaches.

Figure 20 lists the incident in which the user credentials belonging to the demo email id were breached and exposed to the public at large.

Breaches you were pwned in

A "breach" is an incident where data has been unintentionally exposed to the public. Using the 1Password password manager helps you ensure all your passwords are strong and unique such that a breach of one service doesn't put your other services at risk.



bigbasket: In October 2020, the Indian grocery platform bigbasket suffered a data breach that exposed over 20 million customer records. The data was originally sold before being leaked publicly in April the following year and included email, IP and physical addresses, names, phone numbers, dates of birth, passwords stored as Django(SHA-1) hashes.

Compromised data: Dates of birth, Email addresses, IP addresses, Names, Passwords, Phone numbers, Physical addresses



Canva: In May 2019, the graphic design tool website Canva suffered a data breach that impacted 137 million subscribers. The exposed data included email addresses, usernames, names, cities of residence and passwords stored as bcrypt hashes for users not using social logins. The data was provided to HIBP by a source who requested it be attributed to "JimScott.Sec@protonmail.com".

Compromised data: Email addresses, Geographic locations, Names, Passwords, Usernames



Digimon (spam list): In September 2016, over 16GB of logs from a service indicated to be digimon.co.in were obtained, most likely from an unprotected MongoDB instance. The service ceased running shortly afterwards and no information remains about the precise nature of it. Based on enquiries made via Twitter, it appears to have been a mail service possibly based on PowerMTA and used for delivering spam. The logs contained information including 7.7M unique email recipients (names and addresses), mail server IP addresses, email subjects and tracking information including mail opens and clicks.

Compromised data: Email addresses, Email messages, IP addresses, Names



IIMJobs: In December 2018, the Indian job portal IIMJobs suffered a data breach that exposed 4.1 million unique email addresses. The data also included names, phone numbers, geographic locations, dates of birth, job titles, job applications and cover letters plus passwords stored as unsalted MD5 hashes. The data was provided to HIBP by dehashed.com.

Compromised data: Dates of birth, Email addresses, Geographic locations, IP addresses, Job applications, Job titles, Names, Passwords, Phone numbers

Figure 20: List of breaches in which the demo email id was found.

Intelligence X, DeHashed, WikiLeaks and DDO secrets are additional tools that can help uncover information on the surface web as well as on the dark web.

SEARCH ENGINES

Even though Google is the most popular search engine used in India, there are several other alternatives to conduct your search. It is rightly said "*All information is available online; you just need to know how to search for it*".

DUCKDUCKGO

DuckDuckGo is an anonymous search engine that protects user's privacy and avoids the filter of personalized search results like Google.

CARROT2

Carrot 2 is a clustering search engine that collates information across multiple public search engines and creates pie-chart visualization of the clusters.

TOR BROWSER

TOR browser is used to navigate deeper into the dark web. It works for sites with an. onion extension. Other public search engines include Bing, Yahoo and Yandex.

WEBSITE INVESTIGATION

Websites can be vital sources of evidence to gain insight about the enterprise, the people involved and its activities.

DOMAIN SEARCH

www.whois.com allows investigators to trace the ownership of a particular domain name. It maintains a record of information about every domain name purchased along with the relevant dates and its expiry (*figures 21-23*).

facebook.com

Updated 18 hours ago 

Domain Information	
Domain:	facebook.com
Registrar:	RegistrarSafe, LLC
Registered On:	1997-03-29
Expires On:	2028-03-30
Updated On:	2020-03-10
Status:	clientDeleteProhibited clientTransferProhibited clientUpdateProhibited serverDeleteProhibited serverTransferProhibited serverUpdateProhibited
Name Servers:	a.ns.facebook.com b.ns.facebook.com c.ns.facebook.com d.ns.facebook.com

Figure 21: Whois example for domain name facebook.com.

 Registrant Contact	
Name:	Domain Admin
Organization:	Facebook, Inc.
Street:	1601 Willow Rd
City:	Menlo Park
State:	CA
Postal Code:	94025
Country:	US
Phone:	+1.6505434800
Fax:	+1.6505434800
Email:	domain@fb.com

Figure 22: Registrant details for domain name facebook.com.

Administrative Contact	
Name:	Domain Admin
Organization:	Facebook, Inc.
Street:	1601 Willow Rd
City:	Menlo Park
State:	CA
Postal Code:	94025
Country:	US
Phone:	+1.6505434800
Fax:	+1.6505434800
Email:	domain@fb.com

Technical Contact	
Name:	Domain Admin
Organization:	Facebook, Inc.
Street:	1601 Willow Rd
City:	Menlo Park
State:	CA
Postal Code:	94025
Country:	US
Phone:	+1.6505434800
Fax:	+1.6505434800
Email:	domain@fb.com

Figure 23: Whois details about administration and technical contact for domain name facebook.com. Domain names and Internet Protocol (IP) addresses are the framework upon which the entire www (world wide web) is built. Whois IP allows users to lookup IP addresses.

WEB-ARCHIVES

www.waybackmachine.org is an archived directory of digital material ranging from websites, software applications, games, music, videos, movies, and books. The database offers free access to publicly available data, making past websites accessible in the future.

EMAIL INVESTIGATION

Email ids play a crucial role in digital communication systems. All the applications, web portals and digital products require email id as a primary login credential. Every company has a different system of nomenclature for their official email ids. For example, mark.zuckerberg@facebook.com follows the pattern "first name.last name@facebook.com".

With tools like www.email-format.com users can find formats of email ids belonging to thousands of companies worldwide.

www.synapsint.com provides technical information related to email ids. It gauges suspicious and potentially fake emails based on its presence on reputable social media platforms like Twitter and Facebook. Other details like data breach, leaked credentials, domain details etc. are also provided.

www.tools.epioes.com is one of the best email investigation tools for Gmail without notifying the user. The tools runs through social networks and websites like Twitter, Spotify, Snapchat, Samsung, Google, Freelancer, Adobe, Amazon, LinkedIn etc. Furthermore, it returns results related to google reviews, photos that the user has posted on Google, locations that the user has searched amongst other things.

AUTOMATED TOOLS

Gathering information from the internet involves innumerable challenges. Currently, there are about 1.2 million terabytes of data on the internet. Imagine diving into that pool of information and identifying relevant data (Edward J Apple, 2011; Mehta, 2020b). The volatile nature of the internet imposes greater complexities since data can be deleted or manipulated within minutes. The investigations are often carried out in time-pressured environments; however, business related investigations may not involve similar pressure as that of criminal investigations.

One of the essential issues is that OSINT investigators are spending more time in collecting data than analysing it. Automated tools provide faster solutions to data gathering. They use artificial intelligence, machine learning and are able to swim through the deeper and larger portions of the internet. These tools can be useful to large corporates and background verification companies. All these tools allow customizations of reports and offer investigative support. Some of the best tools available are the following:

MALTEGO

Maltego (www.maltego.com) links small pieces of information from different sources and curates advanced visualizations of the search results. It allows users to import, export and pivot through data efficiently (Adel and Cusack, 2020; Pastor-Galindo et al., 2020; US Army, 2012).

SKOPENOW

Skopenow (www.skopenow.com) provides automated intelligence solutions that can turn raw data to actionable intelligence. It scrapes digital footprints of targets and collates relevant information and metadata about them.

COBWEBS

Cobwebs (<https://cobwebs.com>) is an Israeli intelligence product that offers artificial intelligence powered web intelligence solutions to law enforcement agencies globally. The portal integrates features like live data extraction, pattern recognition, deep image analysis, language processing and offers predictive analysis as well.

CONCLUSION

Adopting some of the basic techniques of conducting background verifications can significantly help increase productivity and maintain credibility. Leaked databases may contain sensitive confidential information about the organization; therefore, removing them at the earliest chance possible is both essential and crucial. Hiring employees after conducting thorough background checks will ensure the organization's safety as employees are typically granted immediate access to the organization's IT systems upon joining. , Cyber-attacks may seem off the radar for many businesses, however, there has been a projectile increase in the way trends are emerging today. Every business, nowadays, small or big is prone to such attacks. It is therefore vital that employees and management remain alert to phishing and other forms of cyber-attacks.

Organizations must prioritize cyber awareness programs that will assist employees in conducting their day-to-day activities such as checking emails and finding its source. Website investigations help identify archived pages and activities of the business entity. Verification of legal records ensures safety from fraud, embezzlement, and other crimes. Enterprises may reach faster conclusions to their findings, thereby enabling smoother decision-making capabilities. Journalists also tend to rely on OSINT techniques to fact check submissions/information. Individually owned businesses benefit the most by performing background checks to attract new partnerships, study customer preferences and know their competitors.

Various other techniques like social media investigations (SOCMINT) can also be used to uncover important information. Tools and techniques keep evolving as platforms transform to newer versions owing to privacy measures. However, as an OSINT or background verification specialist, the only predominant factor that can assist with navigation is creativity! Investigation preparation and maps prove to be of greater importance in such a case

REFERENCES

Adel, A. and Cusack, B. (2020) 'Role of Multimedia Information Retrieval in Providing a Credible Evidence for Digital Forensic Investigations: Open Source Intelligence Investigation Analysis', *Computer Science & Information Technology*. AIRCC Publishing Corporation, pp. 11–22. Available at: 10.5121/csit.2020.101002 (Accessed: 20 August 2021).

Bazell, M. (2019) *Open Source Intelligence Techniques*. 7th edn. USA: CreateSpace Independent Publishing Platform.

Button, M. and Cross, C. (2017) *Cyber Frauds, Scams and their Victims*.

Data Report (2021) *Global Social Media Stats — DataReportal — Global Digital Insights*. Data Reportal Available at: <https://datareportal.com/social-media-users> (Accessed: 20 August 2021).

Edward J Apple (2011) *Internet Searches for Vetting, Investigations, and Open-Source Intelligence*. Florida, USA: CRC Press. Available at: https://books.google.com.sg/books?id=hlq6ku99UDYC&pg=PA161&lpg=PA161&dq=osint+investigation+chain+of+custody&source=bl&ots=ZUb4JdREdH&sig=ACfU3U3zQL7fG9UgvQ9uLHbcKt_8ACT-1w&hl=en&sa=X&ved=2ahUKEwje596m2NfnAhUc4zgGHfN2CTQQ6AEwCXoECAcQAQ#v=onepage&q=osint%252 (Accessed: 19 August 2021).

Infosys (2020) Manage in-person KYC, remotely *Infosys Knowledge Institute*. Infosys Knowledge Institute,

Jahankhani, H., Al-Nemrat, A. and Hosseiniyan-Far, A. (2014) 'Cybercrime classification and characteristics', *Cyber Crime and Cyber Terrorism Investigator's Handbook*, (September 2017), pp. 149–164. Available at: 10.1016/B978-0-12-800743-3.00012-8 (Accessed: 18 August 2021).

Mehta, M. (2020a) *Part 7: Internet crimes & Background Vetting for HR using OSINT tools and sites | LinkedIn*, LinkedIn Available at: <https://www.linkedin.com/pulse/part-7-internet-crimes-background-vetting-hr-using-dr-malvika/> (Accessed: 21 August 2021).

Mehta, M. (2020) *Part 1: Introduction to OSINT Investigation (Open Source INTeLLIGENCE) | LinkedIn*, LinkedIn Available at: <https://www.linkedin.com/pulse/introduction-osint-investigation-open-source-dr-malvika/> (Accessed: 21 August 2021).

Pastor-Galindo, J., Nespoli, P., Gomez Marmol, F. and Martinez Perez, G. (2020) 'The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends', *IEEE Access*, 8, pp. 10282–10304. Available at: 10.1109/ACCESS.2020.2965257 (Accessed: 19 August 2021).

Reuser, A.H.P. (2017) 'The RIS Open Source Intelligence Cycle', *Journal of Mediterranean and Balkan Intelligence*, 10(2), pp. 29–44. Available at: <https://arnoreuser.com/wp-content/uploads/2018/12/201712-The-RIS-OSINT-Intelligence-Cycle.pdf> (Accessed: 19 August 2021).

Robert André Furuhaug (2019) *Open Source Intelligence Methodology*. School of Computer Science and Informatics, University College Dublin. Available at: <https://phs.brage.unit.no/phs->

xmlui/handle/11250/2617479 (Accessed: 21 August 2021).

Rosenthal, M. (2020) *Must-Know Phishing Statistics: Updated 2020* | Tessian., Tessian Available at: <https://www.tessian.com/blog/phishing-statistics-2020/#covid-scams-phishing> (Accessed: 18 August 2021).

Samantha Elizabeth Rule (2014) *A Framework on using Open Source Intelligence as a Digital Forensics Investigative Tool*. Rhodes University. Available at: <https://research.ict.ru.ac.za/SNRG/Theses/Rule2014Msc.pdf> (Accessed: 21 August 2021).

Statista (2021a) *Facebook users by country 2021* | Statista., Statista Available at: <https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/> (Accessed: 20 August 2021).

Statista (2021b) *Twitter: most users by country* | Statista., Statista Available at: <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/> (Accessed: 20 August 2021).

The International Consortium of Investigative Journalists (2021) *About* | ICIJ Offshore Leaks Database., Offshore leaks database Available at: <https://offshoreleaks.icij.org/pages/about> (Accessed: 20 August 2021).

US Army (2012) *'Open-Source Intelligence ATP 2-22.9'*, (July) US, p. 91. Available at: <http://www.fas.org/irp/doddir/army/atp2-22-9.pdf> (Accessed: 21 August 2021).